



---

## EDITORIAL

---

### Syndromic Surveillance: a Local Perspective

Farzad Mostashari and Jessica Hartman

A surveillance methodology that was virtually unknown just a few years ago is now poised for deployment across the nation. This is due in no small part to the World Trade Center attacks and the anthrax-laden letters that followed in October 2001. Large-scale bioterrorism now seems likely, if not inevitable, and syndromic surveillance, although largely untested, provides a hope of a precious few hours or days of early warning. But, despite a flood of interest and funding, many questions remain: Just what is “syndromic surveillance”? Which data sources are most promising? Which analytic methods perform best for outbreak detection? How should syndromic surveillance be evaluated? Does it threaten civil liberties and privacy concerns? Is there a legal mandate for it? Who should be conducting the surveillance? Should a national center collect and analyze these data? What is syndromic surveillance good for?

More than 400 public health practitioners, academics, and military and private industry personnel attended the 2002 National Syndromic Surveillance Conference at which these topics were discussed by leading practitioners and theorists, allowing participants to identify key tensions and challenges facing this emerging field. In this editorial, we provide our unvarnished perspective on some of these issues.

#### WHAT IS SYNDROMIC SURVEILLANCE?

Nearly every speaker at the conference challenged the term *syndromic surveillance* as imprecise and potentially misleading. Many of the systems under discussion do not monitor well-defined constellations of signs and symptoms (syndromes), but instead target nonspecific indicators of health, such as a patient with a chief complaint of “cough” or the sale of over-the-counter cold medication. Conversely, many systems that do monitor syndromes (e.g., acute flaccid paralysis, Reye’s syndrome, or carpal tunnel syndrome) are not included in these discussions. Suggested alternative terms include biosurveillance, disease early warning systems, prodromic surveillance, nontraditional surveillance, prediagnostic surveillance, and health indicator surveillance. Because none of these alternatives has generated excitement or

---

Mr. Mostashari is Assistant Commissioner, Bureau of Epidemiological Services, New York City Department of Health and Mental Hygiene, and Ms. Hartman is Research Scientist—Project Director, New York City Department of Health and Mental Hygiene and the New York Academy of Medicine.

Correspondence: Farzad Mostashari, New York City Department of Health and Mental Hygiene, 125 Worth Street, Box CN6, New York, NY 10013. (E-mail: fmostash@health.nyc.gov)

The views and opinions expressed in this editorial are those of the authors and not of the institutions they represent.

gained a constituency, we must continue to use the less-than-ideal label syndromic surveillance.

Two disparate approaches to syndromic surveillance became apparent at the conference. The first aims to detect the incidence of nonspecific mild illness by monitoring data that are routinely collected for other purposes and are available at little or no additional cost (e.g., emergency department visit logs from a hospital billing database). The other approach is to create a data collection system that is dedicated to the purpose of this public health surveillance (e.g., “drop-in” surveillance).<sup>1</sup>

Each approach has advantages and disadvantages. The use of routinely collected data is readily sustainable, benefits from the availability of historic (baseline) data, and imposes no additional burden on busy clinicians. A dedicated data collection system would not be limited by existing information systems, could encompass precise clinical syndromes, and could provide a tool for interacting with and educating physicians. These approaches are sufficiently different from each other that generalizations about one cannot readily be applied to the other. For the purposes of this article, we focus on the more widely used approach, use of routinely collected electronic data for public health surveillance.

### **TRADITIONAL OR NONTRADITIONAL DATA SOURCES**

Some of the more sensational possibilities in syndromic surveillance involve the use of nontraditional data sources, such as tallying orange juice sales at grocery stores or monitoring video cameras to measure the incidence of coughs in public places. Research and development on promising new data sources should continue and should address legitimate privacy concerns.<sup>2</sup>

Most syndromic surveillance systems, however, rely on traditional public health partners like hospitals, emergency departments, and laboratories. Such traditional data sources offer public health officials the ability to “go to the bedside” virtually to obtain more detailed clinical information on the persons involved in the suspected disease cluster, thereby making these data sources the natural starting place for health departments. Increased interactions with traditional partners also provide the possibility of strengthening public health reporting in general.

### **ANALYTIC APPROACHES: TRADITIONAL METHODS, NEW USES**

Andrew Lawson<sup>3</sup> pointed out our lack of statistical techniques specifically designed to detect spatial and temporal outbreaks in real time. Instead, investigators have ingeniously met this need by adapting analytical methods developed for other purposes, such as quality assurance (CUSUM, cumulative sums methods),<sup>4</sup> influenza excess mortality (cyclical regression),<sup>5</sup> and data mining (WSARE, what’s strange about recent events).<sup>6</sup> Funding from the Alfred P. Sloan Foundation has allowed the New York City Department of Health and Mental Hygiene and the New York Academy of Medicine to collaborate with Martin Kulldorff on adapting the methods and software he developed for cancer cluster detection (SaTScan) for use in prospective infectious disease outbreak detection. In this issue, Burkom<sup>7</sup> presents a further adaptation of this method that allows for integration of multiple data streams. To develop and evaluate these methods further, we urgently need to create

semisynthetic validation data sets “spiked” with simulated outbreaks of various shapes and sizes.

### THE CHALLENGE OF EVALUATION

Although no syndromic surveillance system has provided early warning of bioterrorism, these systems are designed to detect outbreaks with hundreds to thousands of potential casualties, and no large-scale bioterrorist attack has occurred since existing systems were instituted. The sensitivity of these systems for detecting such an event can be evaluated by (1) testing the ability of surveillance systems to detect naturally occurring outbreaks, such as annual influenza epidemics and gastrointestinal outbreaks,<sup>5,8–11</sup> or (2) simulating how different large-scale bioterrorist attacks might appear to various “syndromic signal detectors,” as pioneered by Carley<sup>12</sup> and Burkom.<sup>5</sup>

Although it is important to evaluate the sensitivity and predictive value positive of syndromic data for detecting individual cases of target illness, the ability to detect outbreaks in a timely manner is the ultimate objective of these systems. Thus, Dan Sosin of the Centers for Disease Control and Prevention presented a draft framework for evaluating syndromic surveillance systems that emphasizes outbreak detection.<sup>13</sup> A syndromic surveillance system can have high sensitivity and predictive value positive for communitywide outbreaks of respiratory illness,<sup>5</sup> even though the system relies on data that are highly nonspecific and moderately sensitive.<sup>14</sup>

In addition, we need a new evaluation metric that reflects the timeliness of different systems for outbreak detection while simultaneously considering their sensitivity and specificity. Receiver operating characteristic (ROC) curves, such as presented by Burkom,<sup>7</sup> have been used to compare system performance at different threshold levels, but do not value timeliness. Cross-correlation functions<sup>8</sup> can provide a sense of the relative timeliness of different data streams, but are not specifically focused on outbreak detection. Activity monitoring operating characteristic (AMOC) curves,<sup>15</sup> a new methodology applied to financial transaction systems that differentially value signals based on timeliness, on the other hand, show promise for use in syndromic surveillance.

### INVESTIGATION OF SIGNALS

As Don Weiss of the New York City Department of Health and Mental Hygiene remarked, “Everything that has been discussed so far [data sources, analytic methods, evaluation] . . . is the easy part.”<sup>16</sup> Responding to syndromic surveillance signals appropriately may be our most difficult task. Syndromic signals are nondiagnostic, yet if we are to detect a biological attack quickly, we must immediately and effectively investigate warning signals and enhance the diagnostic examination of individuals who otherwise would not have been tested. An emergency room physician at the conference objected to syndromic surveillance systems that ignore front-line clinical staff. An alert clinician or laboratory worker will always be the one to make the diagnosis; our goal is to foster and enhance that alertness.

At the same time, any given signal is most likely not bioterrorism, and we must not exhaust our public health resources and credibility by raising too frequent alarms. In this issue, Duchin and Pavlin<sup>17,18</sup> discuss how their institutions maintain this delicate balance. Clearly, alarm thresholds must be set at a level that is sustainable for public health and reflects the possibility of a bioterrorist attack. Obtaining

appropriate clinical specimens in signal investigations is time consuming and difficult. To maximize our ability to diagnose agents of biologic terrorism rapidly without the disruption of false-positive specimens, it is imperative that we develop sensitive and highly specific rapid diagnostic techniques to facilitate determining the etiology of clusters detected by syndromic surveillance systems.

### **SECURITY VERSUS LIBERTY**

Must civil liberties be compromised in our pursuit of security? A dense and interconnected national database of health and other information, where identifiable and deidentified data are correlated and analyzed, would offer a significant potential for actual or perceived violations of individual privacy. Some ways to lessen this risk include the following:

- Collecting the minimum amount of identifiable data necessary (e.g., age rather than date of birth)
- Assigning data collection and storage to agencies with a legal mandate and tradition of maintaining the confidentiality of individual data, rather than to commercial entities
- Collecting aggregated (count) data rather than individual (line list) data as pursued by the National Bioterrorism Demonstration Project<sup>19</sup>
- Deidentifying or aggregating data used for nonbioterrorism purposes and paying close attention to electronic security<sup>20</sup>
- Using decentralized data networks rather than “warehousing” data at a single, potentially vulnerable, location

### **NATIONAL VERSUS LOCAL**

Some syndromic surveillance data sources are national in their scope, yet can provide information on the local level (e.g., large retail chains, health plans, or health care networks). These data sources offer the promise of achieving broad coverage in a short order at a low cost. Significant limitations to a single national syndromic surveillance system, however, include

- Creating a national analysis center does not compensate for a lack of local capacity and infrastructure. Syndromic surveillance systems are at their core “smoke detectors” and call for immediate investigation and response if they are to provide early warning of outbreaks.
- Although state and local health departments have a legal mandate to collect and analyze patient information for public health surveillance,<sup>21</sup> this mandate does not extend to federal agencies, nonprofit or academic groups, or commercial establishments.
- Centralized analysis of data from hundreds of localities would lead to numerous daily alarms triggered by chance alone. Without knowledge of local events and patterns that might affect the behavior of health systems, patients, and consumers, these systems would run the risk of becoming irrelevant by repeatedly raising false alarms.
- National systems would not have the flexibility to meet rapidly changing, broader local public health needs. For example, after the increase in the ciga-

rette tax in New York City, public health officials were quickly able to confirm an increase in prescriptions for nicotine replacement therapy. Such local dual use would be difficult to implement in national systems.

- Finally, and perhaps most important, a national system would create a single point of failure for bioterrorism surveillance, which would be vulnerable to cyberattacks and physical disruption. In addition, syndromic surveillance is a new field; data sources, syndrome groupings, and analytic methods for outbreak detection have not been standardized, and best practices have not been established. A single point of failure would also occur if the national system chose the wrong data sources or analytic methods.

It make little sense for local health departments to approach large national data providers independently, but any national system with access to such sources should provide local jurisdictions with the data they need to conduct local analyses and the ability to determine response thresholds. Furthermore, national surveillance systems must carefully consider what sorts of timely data and querying systems health departments will need to investigate alarms.

#### **WHO OWNS SYNDROMIC SURVEILLANCE?**

Many entities, including public health departments, academic centers, military contractors, and commercial vendors, can legitimately stake a claim to syndromic surveillance. All share a sense of urgency and civic responsibility, but tensions exist regarding which agency should be the central player in these activities. State and local health departments have the mandate and experience to perform surveillance and investigate disease outbreaks. However, despite the recent infusion of federal resources, health departments face acute shortages in staff trained in information technology, statistics, and computational mathematics. The uncertain and short-term nature of federal bioterrorism funding hinders the ability of health departments to recruit, train, and retain skilled public health personnel. Meanwhile, the academic centers, military contractors, and vendors that can rapidly mobilize the resources necessary to develop various components of syndromic surveillance systems may have little practical experience in outbreak investigations and infectious disease surveillance. One potential model for collaboration would be for public health departments to focus on public health practice using relatively straightforward syndromic surveillance systems based on traditional data sources while other entities pursue research and development and thoroughly evaluate potential data sources, software, and algorithms prior to implementation.

#### **BIOTERRORISM VERSUS EVERYTHING ELSE (DUAL USE)**

Many commentators in public health have questioned the recent increase in funding for bioterrorism-related research, arguing that it diverts resources from already strained public health programs.<sup>22</sup> By not limiting its use to bioterrorism, however, syndromic surveillance can be a boon rather than a drain on traditional public health activities. In his closing remarks at the conference, New York City Commissioner of Health and Mental Hygiene Thomas Frieden emphasized the importance of using syndromic surveillance systems to strengthen all realms of public health. Such “dual uses” in New York City have included

- Determining the epidemiology of drug overdoses
- Monitoring for suicidal ideation following the 1-year anniversary of the World Trade Center attacks
- Quantifying heat-related morbidity
- Monitoring for respiratory distress following insecticide spraying for West Nile virus
- Correlating the decline in asthma hospitalizations to increased use of suppressive therapy for asthma
- Early detection of a citywide gastrointestinal (norovirus) outbreak<sup>20</sup>

## CONCLUSIONS

The promise of syndromic surveillance extends beyond early warning for bioterrorist attacks. Even if bioterrorism is first detected by an astute clinician, syndromic surveillance can help delineate the size, location, and tempo of the epidemic or provide reassurance that a large outbreak is not occurring when a single case or a small, localized cluster of an unusual illness is detected. More broadly, however, as public health and medicine proceed in our information age, the use of existing electronic data for public health surveillance will not appear to be an untested experiment for long. The challenge is to allow these systems to flower without burdening them with unrealistic expectations, centralized control, and unbalanced funding. To help syndromic surveillance systems reach their full potential, we need data standards, guidance to the developers of clinical information systems that will ensure data flow and interoperability, evaluations of best practices, links to improved laboratory diagnostics, regulations that protect privacy and data security, and reliable sustained funding for public health infrastructure to ensure the capacity to respond when the alarm sounds.

## REFERENCES

1. Das D, Weiss D, Mostashari F, et al. Enhanced drop-in syndromic surveillance in New York City following September 11, 2001. *J Urban Health*. 2003;80(2, suppl 1):i76–i88.
2. Newton E. Video and Click-Stream Data as Surveillance Tools. Presented at: The First National Syndromic Surveillance Conference, September 23–24, 2002; New York, NY.
3. Lawson A. Temporal-Spatial Methodologies and Data Sources. Presented at: The First National Syndromic Surveillance Conference, September 23–24, 2002; New York, NY.
4. Hutwagner L, Thompson W, Seeman GM, Treadwell T. The bioterrorism preparedness and response Early Aberration Reporting Systems (EARS). *J Urban Health*. 2003;80(2, suppl 1):i89–i96.
5. Mostashari F, Fine A, Das D, Adams J, Layton M. Use of ambulance dispatch data as an early warning system for communitywide influenzalike illness, New York City. *J Urban Health*. 2003;80(2, suppl 1):i43–i49.
6. Wong W-K, Moore A, Cooper G, Wagner M. WSARE: what's strange about recent events? *J Urban Health*. 2003;80(2, suppl 1):i66–i75.
7. Burkom HS. Biosurveillance applying scan statistics with multiple, disparate data sources. *Journal of Urban Health*. 2003;80(2, suppl 1):i57–i65.
8. Tsui FC, Wagner MM, Dato V, Chang CC. Value of ICD-9 coded chief complaints for detection of epidemics. *Proc AMIA Symp*. 2001:711–715.
9. Lazarus R, Kleinman K, Dashevsky I, et al. Use of automated ambulatory-care encounter records for detection of acute illness clusters, including potential bioterrorism events. *Emerg Infect Dis*. 2002;8:753–760.
10. Goldenberg A, Shmueli G, Caruana RA, Fienberg SE. Early statistical detection of an-

- thrax outbreaks by tracking over-the-counter medication sales. *Proc Natl Acad Sci USA*. 2002;99:5237–5240.
11. Lombardo J, Burkom H, Elbert E, et al. A systems overview of the Electronic Surveillance System for the Early Notification of Community-Based Epidemics (ESSENCE II). *J Urban Health*. 2003;80(2, suppl 1):i32–i42.
  12. Carley K, Yahja A, Kaminsky B. BOWAR city-scale multi-agent network model of bio-attacks. Available at: [www.casos.ece.cmu.edu/projects/BioWar/BioWarPoster.ppt](http://www.casos.ece.cmu.edu/projects/BioWar/BioWarPoster.ppt). Accessed January 31, 2003.
  13. Sosin DM. Draft framework for evaluating syndromic surveillance systems. *J Urban Health*. 2003;80(2, suppl 1):i8–i13.
  14. Greenko J, Mostashari F, Fine A, Layton M. Clinical evaluation of the emergency medical services (EMS) ambulance dispatch-based syndromic surveillance system, New York City. *J Urban Health*. 2003;80(2, suppl 1):i50–i56.
  15. Fawcett T, Provost F. Activity monitoring: noticing interesting changes in behavior. Paper presented at: Fifth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining; August 15–18, 1999; San Diego, CA.
  16. Weiss D. The Public Health Response to Syndromic Surveillance Signals. Presented at: The First National Syndromic Surveillance Conference; September 23–24, 2002; New York, NY.
  17. Duchin JS. Epidemiological response to syndromic surveillance signals. *J Urban Health*. 2003;80(2, suppl 1):i115–i116.
  18. Pavlin JA. Investigation of disease outbreaks detected by “syndromic” surveillance systems. *J Urban Health*. 2003;80(2, suppl 1):i107–i114.
  19. Platt R, Bocchino C, Caldwell B, et al. Syndromic surveillance using minimum transfer of identifiable data: the example of the National Bioterrorism Syndromic Surveillance Demonstration Program. *J Urban Health*. 2003;80(2, suppl 1):i25–i31.
  20. Lober WB, Trigg L, Karras B, et al. Syndromic surveillance using automated collection of computerized discharge diagnoses. *J Urban Health*. 2003;80(2, suppl 1):i97–i106.
  21. Broome CV, Horton HH, Tress D, Lucido SJ, Koo D. Statutory basis for public health reporting beyond specific diseases. *J Urban Health*. 2003;80(2, suppl 1):i14–i22.
  22. Sidel VW, Gould RM, Cohen HW. Bioterrorism preparedness: cooptation of public health? *Med Global Surviv*. 2002;7:82–89.
  23. Norovirus activity—United States, 2002. *MMWR Morb Mortal Wkly Rep*. 2003;52:41–44.